

The neural basis of conceptualizing the same action at different levels of abstraction

Robert P. Spunt,¹ David Kemmerer,² and Ralph Adolphs¹

¹California Institute of Technology, 1200 E. California Blvd., Pasadena, CA, USA and ²Purdue University, West Lafayette, IN, USA

Correspondence should be addressed to Robert P. Spunt, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125 USA. E-mail: spunt@caltech.edu

Abstract

People can conceptualize the same action (e.g. ‘riding a bike’) at different levels of abstraction (LOA), where higher LOAs specify the abstract motives that explain *why* the action is performed (e.g. ‘getting exercise’), while lower LOAs specify the concrete steps that indicate *how* the action is performed (e.g. ‘gripping handlebars’). Prior neuroimaging studies have shown that *why* and *how* questions about actions differentially activate two cortical networks associated with mental-state reasoning and action representation, respectively; however, it remains unknown whether this is due to the differential demands of the questions *per se* or to the shifts in LOA those questions produce. We conducted functional magnetic resonance imaging while participants judged pairs of action phrases that varied in LOA and that could be framed either as a *why* question (Why ride a bike? Get exercise.) or a *how* question (How to get exercise? Ride a bike.). Question framing (*why* vs *how*) had no effect on activity in regions of the two networks. Instead, these regions uniquely tracked parametric variation in LOA, both across and within trials. This suggests that the human capacity to understand actions at different LOA is based in the relative activity of two cortical networks.

Key words: action understanding; semantic memory; abstraction; concepts; social cognition; fMRI

Introduction

People have a rich set of concepts for representing actions and a flexible capacity to use these concepts to understand both their own and other people’s actions. This allows humans to not only understand the meaning of different actions, but also understand that the *same* action can carry different meanings. For instance, an action initially identified as ‘writing a manuscript’ might be further conceptualized in terms of its abstract motives—i.e. *why* it is done (e.g. ‘sharing knowledge’)—or its concrete implementation—i.e. *how* it is done (e.g. ‘typing words’). The present study attempted to clarify the neural basis of conceptualizing an action at varying levels of abstraction (LOA).

Several recent neuroimaging studies have shown that answering *why* vs *how* questions about action reliably differentiates activity in two left-lateralized cortical networks (Spunt et al., 2010, 2011; Spunt and Lieberman, 2012a,b; Spunt and Adolphs, 2014). Specifically, the *Why* > *How* contrast reliably reveals activation in the dorsomedial and ventromedial

prefrontal cortex (PFC), the anterior superior temporal sulcus (STS), the temporoparietal junction (TPJ) and the posterior cingulate cortex (PCC)—regions that have been independently implicated in representing and reasoning about the mental states that typically drive actions, such as beliefs, desires and intentions (Gallagher and Frith, 2003; Saxe, 2006; Carrington and Bailey, 2009; Van Overwalle and Baetens, 2009; Lieberman, 2010; Mar, 2011; Denny et al., 2012; Schurz et al., 2014). Conversely, the *How* > *Why* contrast reliably activates the dorsal and ventral premotor cortex (PMC), posterior middle temporal gyrus (MTG), rostral inferior parietal lobule (IPL) and dorsal precuneus—regions that have been independently implicated in representing the visual motion patterns and somatomotor features of actions when perceived and performed (Caspers et al., 2010; Molenberghs et al., 2012; Rizzolatti et al., 2014), conceptualized (Kemmerer et al., 2012; Watson et al., 2013; Urgesi et al., 2014), and verbally processed (Kemmerer et al., 2012; Pulvermuller, 2013; Kemmerer, 2015).

Received: 20 March 2015; Revised: 19 May 2015; Accepted: 24 June 2015

© The Author (2015). Published by Oxford University Press. For Permissions, please email: journals.permissions@oup.com

The reliable cortical dissociation of *why* and *how* questions is based on an attentional manipulation referred to as the Why/How Task (Freitas et al., 2004), where participants are asked to answer (or evaluate answers to) *why* and *how* questions about the same stimuli (e.g. named or visually presented actions). This ensures that any observed effects are caused by the different cognitive demands of answering a *why* vs *how* question about the same action. Yet, the nature of these differential demands—and their relationship to the *why/how* distinction—remains unclear. Here, we evaluate two alternative accounts of the robust neural effects observed in the *why/how* contrast. These alternatives are suggested by the hierarchical structure of human representations of action (Vallacher and Wegner, 1987; Kozak et al., 2006; see also Fujita et al., 2006; Trope and Liberman, 2010). To illustrate, consider the role that *why* and *how* questions play in navigating the levels of the conceptual action hierarchy in Figure 1A. On the one hand, *why* and *how* questions elicit directional (signed) shifts up and down the levels of the hierarchy, respectively. Taking as examples the intermediate-level phrases ‘make a phone call’ and ‘contact a friend’, a *why* question enables an upward movement (‘*Why* make a phone call? Contact a friend’), while a *how* question enables the reverse (‘*How* to contact a friend? Make a phone call’). This suggests that the cortical networks modulated by the *why/how* contrast may underlie the distinct ‘mindsets’ (Gollwitzer et al., 1990) for thinking about actions in terms of their motive vs their implementation.

This account is complicated by the fact that the language used to represent an action becomes progressively less concrete, specific, imageable, embodied and emotionally neutral as one ascends the levels of the hierarchy. Henceforth, we use the general

term LOA to refer to this basic semantic dimension correlated with increasing levels of a conceptual action hierarchy. Because previous neuroimaging studies contrasted *why* and *how* questions about the same actions (e.g. ‘*Why/How* to contact a friend?’), the answers yielded by *why* questions were likely conceptually higher in LOA (e.g. ‘Feel connected’) than those yielded by corresponding *how* questions (e.g. ‘Make a phone call’). For this reason, the effective ingredient in the *why/how* contrast may be the differential demands of conceptualizing the same action at relatively higher vs lower LOAs. In the present study, we designed a novel Why/How Task that separates the *why/how* question manipulation from the changing conceptual demands imposed by understanding actions at different LOAs. This task design exploits the fact that *why* and *how* questions can be posed at any level of a conceptual action hierarchy, making it possible for a *why* question (‘*Why* dial numbers?’) to yield an answer that is less abstract (‘Make a phone call’) than an answer (‘Contact a friend’) to a *how* question (‘*How* to feel connected?’). This orthogonality allowed us to evaluate the alternative interpretations of the *why/how* contrast described above.

Materials and method

Participants

Nineteen right-handed adults participated in the study in exchange for financial compensation. Participants were neurologically and psychiatrically healthy, had normal/corrected-to-normal vision, spoke English fluently, had IQ in the normal range (assessed with the Wechsler Abbreviated Scales of Intelligence), and were not pregnant. Participants provided

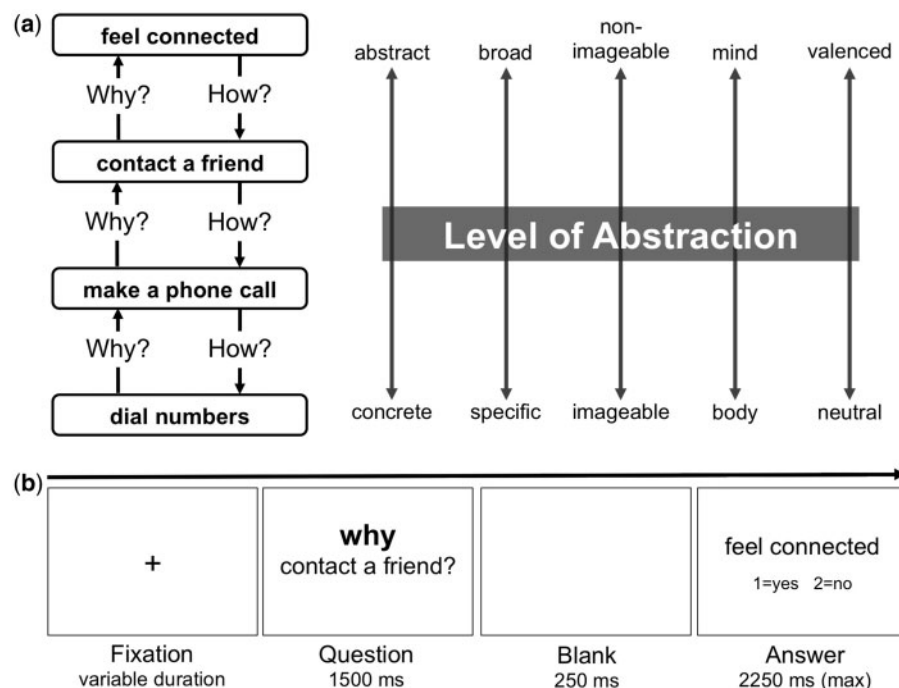


Fig. 1. Experimental design. (A) One of 25 four-level action hierarchies featured in the experimental task. For each four-level hierarchy the action described by a phrase at one level (e.g. ‘make a phone call’) was both a commonly accepted motive for performing the action described by the phrase at the level immediately below it (e.g. ‘dial numbers’) and a commonly accepted means for performing the action described by the phrase at the level immediately above it (e.g. ‘contact a friend’). As shown to the right of the example and described further in the text all phrases were normed on five dimensions used to derive a single factor describing each phrase’s level of abstraction (LOA). (B) Structure of one of the trials formed by pairing phrases at contiguous levels of the hierarchy. Trials began with an action phrase embedded in either a *why* (shown) or *how* question and concluded with a different action phrase presented as a possible answer. Answer phrases were presented for a maximum duration of 2250 ms. Once the participant responded, the screen was replaced by a fixation cross until the onset of the next trial.

written informed consent according to a protocol approved by the Institutional Review Board of the California Institute of Technology. Data from two participants were excluded due to poor task performance (no response to greater than 10% of trials). This left 17 participants for the analysis (11 males, 6 females; mean age = 29.47, age range = 21–46). A power analysis of the Why/How contrast from Spunt and Adolphs (2014) indicated the present study's sample size was sufficiently large to test our hypotheses (see Supplementary Methods).

Action stimuli

Stimuli consisted of 128 syntactically similar phrases, each of which used a simple verb–complement construction (typically a transitive verb and its direct object) to describe a familiar type of human action (the complete set is provided in Supplementary Table S1). As illustrated in Figure 1A, each phrase was part of a grouping of four phrases that collectively formed a conceptual action hierarchy, such that the action described by a phrase at one level (e.g. 'Make a phone call') was both a commonly accepted motive for the action described by a phrase at the level below it (e.g. 'Dial numbers'), and a commonly accepted means for the action described by the phrase at the level above it (e.g. 'Contact a friend'). By pairing phrases at contiguous levels, each set of four phrases yielded three *why* question–answer pairs (e.g. 'Why dial numbers? Make a phone call') and three *how* question–answer pairs (e.g. 'How to make a phone call? Dial numbers').

The final group of 25 question–answer pairs was selected by identifying those that elicited the highest degree of yes/no response agreement in an independent sample of approximately 40 native English-speaking adults recruited online via Amazon.com's Mechanical Turk (see Supplementary Appendix A for details). This resulted in 150 total test trials, 30 of which were foils in that they featured question–answer pairs that were commonly rejected (e.g. 'How to show ambition? Serve alcohol'). These foils were included to check task comprehension and were not of interest in the fMRI analyses.

We characterized each verb–complement phrase (without the *why* and *how* question markers) on several lexical dimensions, including the number of characters and words and the average frequency of content words (per million words in the SUBTLEX database). In addition, we used independent groups of native English-speaking Mechanical Turk participants to norm each phrase on five semantic dimensions believed to covary with the LOA (see Supplementary Appendix A for details): (i)

abstract vs concrete ($N = 150$), (ii) non-imageable vs imageable ($N = 97$), (iii) broad vs specific ($N = 97$), (iv) mind-dependent vs body-dependent ($N = 132$) and (v) emotionally valenced vs neutral ($N = 96$), which was calculated as the absolute deviation from the neutral point on a bipolar (negative/positive) scale. Inter-rater reliability for all measures was excellent (minimum intra-class correlation coefficient = 0.98; Shrout and Fleiss, 1979). Given their substantial inter-correlation (Table 1), we subjected the five semantic dimensions to a principal component analysis (PCA) using the MATLAB Statistics Toolbox (version R2014b; MathWorks Inc., Natick, MA, USA). The PCA showed that a single component explains 91.01% of their total variance (first eigenvariate = 8.12, second eigenvariate = 0.42). Consequently, we used the scores for this underlying dimension to quantify the LOA for each action phrase.

As shown in Table 2, independent-samples *t*-tests showed that *Why* and *How* trials were matched not only on all lexical parameters, but also on the average LOA of the question and answer in each trial, which we henceforth call 'Trialwise LOA.' As described below, our first two analyses exploit the orthogonality of the *Why/How* manipulation and Trialwise LOA to identify their independent effects. In contrast, our third analysis exploits the fact that *Why* and *How* trials differed in what we henceforth call the 'Signed LOA Shift', which refers to the within-trial change in LOA introduced by the answer phrase relative to the question phrase. This is calculated for each trial by subtracting the LOA of the question phrase (henceforth called the 'Prepotent LOA') from the LOA of the answer phrase. In line with the view that *why* and *how* questions produce relative increases and decreases in LOA, respectively, *Why* trials were associated with positive Signed LOA Shifts while *How* trials were associated with negative Signed LOA Shifts (Table 2).

Experimental task

The 150 test trials were presented to participants in an event-related design. The trial structure and timing are shown in Figure 1B. Trials were arranged in a pseudorandom order with a variable onset asynchrony drawn from a pseudoexponential distribution with a mean of 6000 ms (Min/Max = 5000/10 000 ms). The order and onsets of trials were optimized for estimating the *why/how* contrast. This was achieved by generating the design matrices for one million pseudorandomly generated designs, and for each summing the efficiencies of *Why* > *How* contrast estimation. The most efficient design was used for all participants.

Table 1. Pearson correlations among all phrase-level parameters

Parameter	1	2	3	4	5	6	7	8	9	10
1 Level of abstraction	-	0.99	0.97	0.96	0.92	0.77	0.16	-0.25	-0.03	0.14
2 Abstract/concrete	0.99	-	0.96	0.95	0.89	0.74	0.17	-0.26	-0.06	0.14
3 Non-Imageable/imageable	0.97	0.96	-	0.93	0.83	0.76	0.14	-0.24	-0.07	0.15
4 Broad/specific	0.96	0.95	0.93	-	0.82	0.74	0.16	-0.28	-0.07	0.13
5 Mind/body	0.92	0.89	0.83	0.82	-	0.64	0.14	-0.15	0.05	0.1
6 Valenced/neutral	0.77	0.74	0.76	0.74	0.64	-	0.13	-0.2	0.13	0.1
7 Number of characters	0.16	0.17	0.14	0.16	0.14	0.13	-	0.04	-0.35	0.02
8 Number of words	-0.25	-0.26	-0.24	-0.28	-0.15	-0.2	0.04	-	0.45	0
9 Content word frequency	-0.03	-0.06	-0.07	-0.07	0.05	0.13	-0.35	0.45	-	-0.06
10 Response time (s)	0.14	0.14	0.15	0.13	0.1	0.1	0.02	0	-0.06	-

Notes. The first parameter—level of abstraction—represents the first eigenvariate from a principal components analysis on parameters 2–6. Parameters 7–10 were included as parametric nuisance covariates in all single-subject fMRI analyses (represented at the trial level by averaging the parameter value for the question and answer phrases). Note that because RT is a trial-level parameter and is contingent on participant behavior, all RT correlation values were computed as the mean within-subject correlation between RT to each trial and the average parameter value across the question and answer phrases within each trial.

Table 2. Means for performance, lexical and level of abstraction (LOA) parameters for both Why and How trials

Parameter group	Why	How	P _{difference}
Parameter name			
Performance			
Percent correct	94.72	94.14	0.463
Response time (s)	1.00	1.02	0.296
Lexical			
Number of characters	11.90	12.02	0.712
Number of words	2.46	2.42	0.700
Content word frequency	3.80	3.81	0.957
Level of abstraction			
Question-and-answer	0.44	0.44	0.964
Question	0.34	0.59	<0.001
Answer	0.57	0.32	<0.001
Answer-question	0.23	-0.27	<0.001

Notes. The final column lists the P value from a t-test comparing the Why and How trials on each parameter. Paired samples t-tests were used to compare the performance parameters while independent samples t-tests were used for the remaining parameters. To facilitate comparability, the LOA parameters were computed after rescaling the LOA across phrases to 0–1.

Stimulus presentation and response recording were implemented with the MATLAB Psychophysics Toolbox (version 3.0.9; Brainard, 1997). An LCD projector was used to present the task on a screen at the rear of the scanner bore that was visible to participants through a mirror positioned on the head coil. Participants were given a button box and made their yes/no responses with their right-hand index/middle fingers. Prior to the experimental task, participants performed a practice version featuring trials not used in the experimental task.

Image acquisition

All imaging data were acquired at the Caltech Brain Imaging Center using a Siemens Trio 3.0 Tesla MRI scanner outfitted with a 32-channel phased-array headcoil. For the experimental task, we acquired 909 T2*-weighted echoplanar image volumes (EPIs; multiband acceleration factor = 4, slice thickness = 2.5 mm, in-plane resolution = 2.5×2.5 mm, 56 slices, Repetition Time = 1000 ms, Echo Time = 30 ms, flip angle = 60° , Field-of-View = 200 mm). Participants' in-scan head motion was minimal (max translation = 1.93 mm, max rotation = 1.75°). We also acquired an additional 1330 EPI volumes for each participant as part of a separate study. Finally, we acquired a high-resolution anatomical T1-weighted image (1 mm isotropic) and fieldmaps used to estimate and correct for inhomogeneity-induced image distortion.

Image preprocessing

Images were analyzed using Statistical Parametric Mapping (SPM8, Wellcome Department of Cognitive Neurology, London, UK). Each participant's EPI timeseries was subjected to the following preprocessing steps: the first four volumes were discarded to account for T1-equilibration effects; the realign and unwarp procedure was used to perform distortion correction and motion correction; their T1 structural volume was co-registered to the mean of the corrected EPI volumes; group-wise DARTEL registration (Ashburner, 2007) was used to normalize the T1 volume to a group-specific template, with subsequent affine registration to Montreal Neurological Institute (MNI) space; and all EPI volumes were normalized to MNI space using the

deformation flow fields from the previous step, which re-sampled volumes ($2 \times 2 \times 2$ mm) and applied spatial smoothing (6 mm Gaussian kernel, full-width at half-maximum).

Single-subject regression models

General linear models were used to estimate three models of the EPI timeseries for each participant. The following procedures were used in all three models. First, covariates of interest excluded foil trials and trials to which the participant gave either no response or the incorrect response. Excluded trials were modelled in a separate nuisance covariate. Second, the neural response to each trial was defined with variable epochs (Grinband et al., 2008) spanning the onset of the question phrase and the offset of the answer phrase (see Figure 1B). Third, all models included trialwise nuisance covariates for variability in response time (RT), character count, word count and average content word frequency. The three lexical parameters pertain only to the simple verb-complement action phrases, without the *why* and *how* question markers. Each of these nuisance covariates was constructed by modulating the amplitude of the predicted neural response for each trial of interest by the de-meaned parameter value. Fourth, all models included scan-wise nuisance covariates for the six motion parameters estimated from image realignment and a predictor for every timepoint where in-brain global signal change (GSC) exceeded 2.5 SDs of the mean GSC or where estimated motion exceed 0.5 mm of translation or 0.5° of rotation. Finally, all models used the canonical (double-gamma) response function and its temporal derivative to model the hemodynamic response; were high-pass filtered at 1/128 Hz; and were estimated using the SPM8 RobustWLS toolbox, which implements robust weighted least-squares estimation (Diedrichsen and Shadmehr, 2005).

Estimating the Independent Effects of Why/How Questions and Trialwise LOA. The first two models were designed to assess the independent contributions of two factors: variation in the question in each trial (Why vs How), and variation in the Trialwise LOA (average LOA for question and answer phrases). The models differed only in their representation of the factor corresponding to Trial-wise LOA.

The first model included four covariates of interest corresponding to the cells created by crossing factors corresponding to the Question (Why vs How) and Trialwise LOA binarized into a two-level factor (High vs Low). We binarized LOA using Otsu's method (implemented using GRAYTHRESH in MATLAB), which selects the threshold that minimizes the variance within the two bins. Independent-samples t-tests confirmed that Trialwise LOA for trials binned as High LOA was significantly higher than that for trials binned as Low LOA, $t(109) = 20.57$, $P < 0.001$. Following estimation, we computed the [(High-Why + Low-Why) > (High-How + Low-How)] contrast to identify regions independently modulated by the Why/How factor, and the [(High-Why + High-How) > (Low-Why + Low-How)] contrast to identify regions independently modulated by Trialwise LOA.

The second model included two covariates of interest corresponding to trialwise variability in the Question (Why vs How) and in LOA. These covariates were constructed by modulating the height of the predicted neural response to each trial by a value representing the Question (Why = +1; How = -1) and Trialwise LOA. In addition to the nuisance covariates specified above, this model also included a covariate corresponding to the fixed-amplitude (time-invariant) response to each trial. For subsequent group analysis, we computed two contrast images,

one for the modulator coding the Why/How contrast, and one for the modulator coding Trialwise LOA.

Estimating the independent effects of the prepotent LOA and signed LOA shift. The third and final model exploits the fact that although Why and How trials were equated on Trialwise LOA, they differed in the direction of the Signed LOA Shift (Table 2). Moreover, the magnitude of these shifts varied across trials in a way uncorrelated with variation in Trialwise LOA ($r = -0.02$). Figure 4A shows that the magnitude of such shifts showed a moderate negative correlation with the Prepotent LOA ($r = -0.49$), defined earlier as the LOA of the question phrase. If the active ingredients of *why/how* questions are *increases/decreases* in LOA, respectively, then brain regions associated with the *why/how* contrast in prior work should track the magnitude of the Within-Trial Shift in LOA relative to the Prepotent LOA. Hence, this model included two covariates of interest corresponding to trialwise variability in both of these measures. To ensure that any shift-related effects were due to *signed* shifts, we included an additional nuisance covariate for the *unsigned* (absolute) LOA shifts. As in the second model, this model also included a covariate of no interest corresponding to the fixed-amplitude response to each trial. For subsequent group analysis, we computed two contrast images, one for the modulator coding the Prepotent LOA, and one for the modulator coding the Signed LOA Shift.

Group analyses

Contrasts were first interrogated using a set of independently defined left-hemisphere regions of interest (ROI) based on the group-level Why/How contrasts from Study 1 and Study 3 reported in Spunt and Adolphs (2014). These images are publicly available on NeuroVault (<http://neurovault.org/collections/445/>). We selected five ROIs from both the Why > How and How > Why contrasts (Figure 2). The peak coordinate and spatial extent of each ROI is provided in Supplementary Table S2. For each ROI, we tested our hypotheses with *t*-tests on the average parameter estimate across voxels. Confidence intervals (CIs) were estimated using the bias-corrected and accelerated percentile method (10,000 random samples with replacement; implemented using BOOTCI in MATLAB).

ROI analyses were complemented by whole-brain analyses. We conducted one-sample *t*-tests on single-subject contrast images for effects of interest. The resulting group-level *t*-statistic images were interrogated by applying a cluster-forming threshold of $P < 0.001$ followed by cluster-level correction at a family-wise error (FWE) of 0.05. Thresholded results were surface rendered using SurfPlot (<http://mrtools.mgh.harvard.edu/index.php/SurfPlot>).

Results

Behavioral results

Table 2 displays the mean percentage correct and RT to correct responses across Why and How trials. Paired-sample *t*-tests yielded no evidence for Why/How effect on either outcome ($P_s > 0.295$). Given that the LOA factors examined in the regression models were based only on trials with correct responses, we tested the effect of LOA only on RT to correct trials. We first tested this with the binarized LOA factor used in the factorial model. A paired-samples *t*-test showed that RTs to High LOA trials ($M = 1035$ ms, $SD = 143$) were longer than RTs to Low LOA

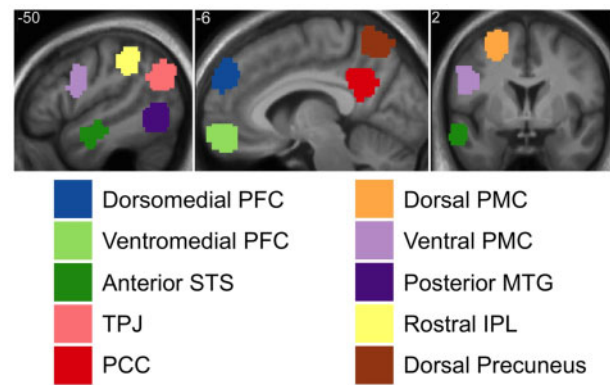


Fig. 2. Left hemisphere cortical regions reliably modulated by the Why/How contrast in prior work. See *Methods* and Supplementary Table S2 for further details. TPJ = Temporoparietal Junction; PFC = Prefrontal Cortex; STS = Superior Temporal Sulcus; PCC = Posterior Cingulate Cortex; IPL = Inferior Parietal Lobule; PMC = Premotor Cortex; MTG = Middle Temporal Gyrus.

trials ($M = 965$ ms, $SD = 137$), $t(16) = 5.52$, $P < 0.001$, 95% CI_{boot} [48, 96]. Next, we tested for a non-zero within-subject correlation of RT and Trialwise LOA. A one-sample *t*-test showed that the within-subject correlation between RT and Trialwise LOA ($r_{mean} = 0.14$, $r_{SD} = 0.08$) was reliably above zero, $t(16) = 6.89$, $P < .001$, 95% CI_{boot} [0.10, 0.17].

These effects of Trialwise LOA on RT underscore the importance of controlling for within-subject performance variability on the measured fMRI signal. Indeed, as shown in Supplementary Table S5 and Figure S1, longer RTs were indeed associated with activity in a distributed cortical network consistent with regions observed in previous studies examining the executive aspects of semantic memory use (Badre et al., 2005; Binder et al., 2005; Goldberg et al., 2007; Hoffman et al., 2010; Raposo et al., 2012; Satpute et al., 2014). Critically, the imaging results presented below are statistically independent of these RT effects.

Imaging results

Effects of why/how questions and trialwise LOA. Table 3 shows ROI-specific results for the first and second models that examine the independent contributions of Question (Why vs How) and binarized Trialwise LOA. Remarkably, neither model produced evidence for a Why/How effect in any of the 10 ROIs examined. In contrast, the Trialwise LOA factor in both models showed a reliable effect on the same 8/10 ROIs, including 5/5 of the ROIs selected for showing an effect in the Why > How contrast in prior studies.

These findings are largely reproduced in the whole-brain analysis (Supplementary Table S3). Neither model yielded evidence for a reliable Why/How effect in any region. In contrast, the Trialwise LOA factor modulated a largely left-lateralized cortical network that was consistently localized across both models. As shown in Figure 3, higher LOAs were consistently associated with regions of the medial PFC, anterior STS, temporal pole, TPJ, and precuneus, while lower LOAs were consistently associated with the posterior MTG, rostral IPL, and regions of the IFG around the pars triangularis. Several clusters were only observed in the second model examining parametric variation in Trialwise LOA. This included bilateral regions of middle occipital cortex that were associated with increasing Trialwise LOA. Although not predicted, this finding is consistent with

Table 3. Results of two-tailed paired-sample *t*-tests on percent signal change in the set of *a priori* regions of interest (ROI) for the two models examining the independent effects of Why/How manipulation and trialwise level of abstraction (LOA)

ROI source Region name	Model 1						Model 2					
	Question: Why > How			LOA: High > Low			Question: Why > How			Increasing LOA		
	<i>t</i>	<i>P</i>	95% CI	<i>T</i>	<i>P</i>	95% CI	<i>t</i>	<i>P</i>	95% CI	<i>t</i>	<i>P</i>	95% CI
<i>Why > How</i>												
Dorsomedial PFC	0.54	0.595	[−0.12, 0.17]	3.84	0.001	[0.21, 0.65]	0.63	0.535	[−0.1, 0.15]	4.00	0.001	[0.04, 0.11]
Ventromedial PFC	0.24	0.816	[−0.11, 0.15]	4.31	<0.001	[0.17, 0.42]	0.57	0.577	[−0.06, 0.18]	3.32	0.004	[0.02, 0.08]
Anterior STS	0.93	0.366	[−0.02, 0.07]	5.67	<0.001	[0.09, 0.19]	1.21	0.242	[−0.02, 0.07]	4.49	<0.001	[0.02, 0.04]
TPJ	0.77	0.454	[−0.06, 0.13]	4.06	<0.001	[0.10, 0.27]	0.41	0.688	[−0.08, 0.12]	2.77	0.014	[0.01, 0.05]
PCC	0.24	0.811	[−0.07, 0.09]	4.01	0.001	[0.08, 0.24]	0.21	0.840	[−0.08, 0.09]	2.95	0.009	[0.01, 0.05]
<i>How > Why</i>												
Dorsal PMC	−0.57	0.575	[−0.09, 0.06]	−1.11	0.283	[−0.16, 0.03]	−1.15	0.269	[−0.11, 0.04]	−1.19	0.252	[−0.06, 0]
Ventral PMC	−0.08	0.939	[−0.07, 0.07]	−2.42	0.028	[−0.27, −0.05]	−0.65	0.526	[−0.09, 0.04]	−2.99	0.009	[−0.07, −0.02]
Posterior MTG	0.05	0.960	[−0.07, 0.08]	−3.57	0.003	[−0.36, −0.12]	−0.35	0.734	[−0.09, 0.06]	−3.99	0.001	[−0.09, −0.03]
Rostral IPL	0.62	0.543	[−0.03, 0.08]	−2.52	0.023	[−0.27, −0.05]	0.24	0.814	[−0.04, 0.07]	−2.65	0.018	[−0.06, −0.01]
Dorsal precuneus	−0.21	0.837	[−0.11, 0.07]	1.57	0.135	[−0.01, 0.2]	−0.62	0.547	[−0.13, 0.06]	1.57	0.136	[−0.01, 0.04]

Note. Both models featured the same two covariates of interest: variation in the question posed in each trial (Why vs How) and trialwise variation in the LOA, computed as the mean LOA for the Question and Answer phrases appearing within the trial. Whereas Model 1 examines the trialwise LOA as a categorical variable with two levels (High vs Low), the Model 2 examines it as a continuous parametric modulator of the response to each trial (see Methods for further details regarding model construction and contrast calculations). The examined ROIs are displayed in Figure 2 and further details can be found in the main text and in Supplementary Table S2. Confidence intervals (CIs) for the effect size in each comparison were estimated using the bias corrected and accelerated percentile method (10000 random samples with replacement; implemented using the BOOTCI function in MATLAB).

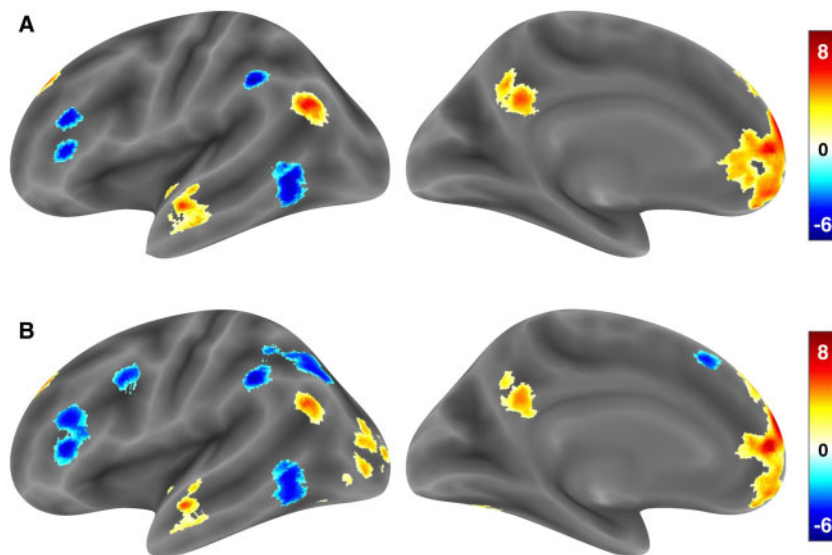


Fig. 3. Whole-brain surface renderings of regional activity associated with Trialwise Level of Abstraction (LOA), computed as the mean of the question and answer phrases appearing in each trial. The results in (A) show regions modulated in the categorical High > Low LOA contrast from Model 1, while the results in (B) show regions modulated by the continuous Trialwise LOA parameter from Model 2. Significant clusters were identified in a whole-brain search thresholded at a cluster-level family-wise error rate of 0.05, and their locations are reported in Supplementary Table S3. To provide information about extent and for display purposes only, the cluster-corrected maps were minimally dilated prior to surface rendering.

recent work showing a left middle occipital association with abstraction (Gilead et al., 2014). Moreover, decreasing Trialwise LOA yielded more extensive regional associations in the second model, including regions of the ventral PMC and presupplementary motor area.

Effects of the prepotent LOA and signed LOA shift. Table 4 shows ROI-specific results for the third model examining the Prepotent LOA and the Signed LOA Shift. Parametric variation in Prepotent LOA showed largely the same effects observed for the

Trialwise LOA parameters examined by the first two models, and this observation was reproduced by the whole-brain analysis (Figure 4B; Supplementary Table S4). More interestingly, however, when examining the Signed LOA Shift, we found evidence for a positive association with four of the five ROIs from the Why > How contrast (the effect in the TPJ ROI was also in the expected direction, $P = 0.062$). And although none of the *a priori* ROIs based on the How > Why contrast showed a significant negative association with the Signed LOA Shift, four of five were in the expected direction. In the whole-brain analysis, the

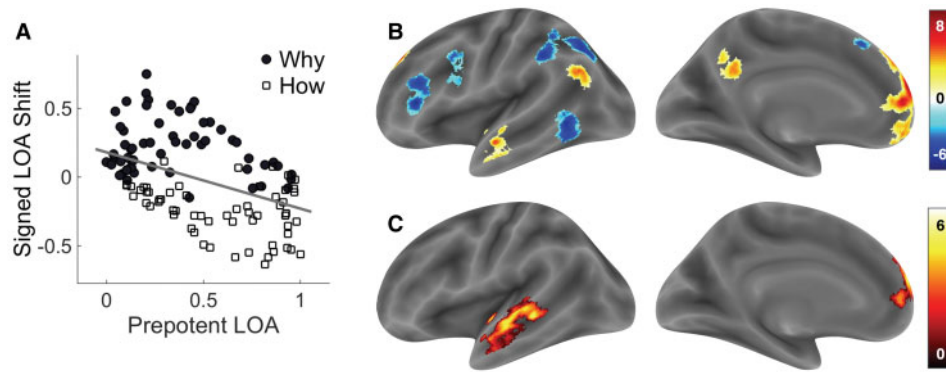


Fig. 4. (A) Scatterplot showing the relationship between the Prepotent LOA and the signed Within-Trial Shift in LOA across Why (black markers) and How (white markers) trials. The Prepotent LOA refers to the LOA of the action phrase appearing in the question at the beginning of each trial. The signed within-trial shift in LOA is computed for each trial by subtracting the Prepotent LOA from the LOA of the action phrase appearing in the presented answer. Positive shift trials induced an upward change in LOA, while negative shift trials induced a downward change in LOA. To facilitate comparability, the LOA dimension has been rescaled to 0–1. (B) Regions uniquely associated with the Prepotent LOA parameter when controlling for the signed within-trial shift in LOA. (C) Regions uniquely associated with the signed within-trial shift in LOA parameter. Significant clusters were identified in a whole-brain search thresholded at a cluster-level family-wise error rate of 0.05, and their locations are reported in Supplementary Table S4. To provide information about extent and for display purposes only, the cluster-corrected maps were minimally dilated prior to surface rendering.

Table 4. Results of two-tailed paired-sample *t*-tests on percent signal change in the set of *a priori* regions of interest (ROI) for the model examining the independent effects of the prepotent level of abstraction (LOA) and the signed shift in LOA for each trial

ROI source	Effect name					
	Prepotent LOA			Signed LOA Shift		
Region name	<i>t</i>	<i>P</i>	95% CI	<i>t</i>	<i>P</i>	95% CI
<i>Why > How</i>						
Dorsomedial PFC	4.53	<0.001	[0.04, 0.10]	2.87	0.011	[0.01, 0.07]
Ventromedial PFC	3.10	0.007	[0.02, 0.08]	2.81	0.013	[0.02, 0.07]
Anterior STS	4.47	<0.001	[0.02, 0.04]	4.02	<0.001	[0.01, 0.03]
TPJ	3.06	0.008	[0.01, 0.06]	2.00	0.062	[0, 0.05]
PCC	2.75	0.014	[0.01, 0.05]	2.63	0.018	[0.01, 0.04]
<i>How > Why</i>						
Dorsal PMC	−1.15	0.266	[−0.06, 0]	−0.96	0.351	[−0.03, 0.01]
Ventral PMC	−2.78	0.014	[−0.06, −0.01]	−1.83	0.086	[−0.05, 0]
Posterior MTG	−4.17	<0.001	[−0.08, −0.03]	−1.88	0.078	[−0.04, 0]
Rostral IPL	−2.70	0.016	[−0.06, −0.01]	−0.95	0.358	[−0.03, 0.01]
Dorsal precuneus	1.66	0.117	[−0.01, 0.05]	1.99	0.064	[0, 0.04]

Note. The model featured two covariates of interest: the LOA of the Question phrase (i.e. Prepotent LOA) and the signed shift in LOA introduced in the Answer phrase, computed by subtracting the LOA of the Question phrase from that of the Answer phrase (see Methods for further details regarding model construction and contrast calculations). The examined ROIs are displayed in Figure 2 and further details can be found in the main text and in Supplementary Table S2. Confidence intervals (CIs) for the effect size in each comparison were estimated using the bias corrected and accelerated percentile method (10 000 random samples with replacement; implemented using the BOOTCI function in MATLAB).

dorsomedial PFC and the anterior STS showed a positive association with the Signed LOA Shift (Figure 4C; Supplementary Table S4). No regions, however, were found to show a negative association with the Signed LOA Shift.

Discussion

We investigated the neural basis of conceptualizing the same actions at different LOAs. This was directly motivated by previous neuroimaging studies showing that *why* and *how* questions about actions differentially activate cortical networks associated with mental-state reasoning and action representation, respectively. Since these studies always asked *why* and *how* questions about the same action stimuli (named or visually presented), they confounded the task manipulation (*why/how* questions) with the use of action concepts that varied in LOA. We

deconfounded these two factors in order to evaluate two alternative functional accounts of the cortical networks known to be modulated by the *why/how* contrast. We found no support for the account that *why* and *how* questions per se elicit distinct and content-free cognitive sets for conceptualizing action. Instead, the evidence supported the alternative account that the distinct effects of *why* and *how* questions are caused by the relative increases and decreases in the conceptual LOA they tend to produce, respectively.

Brain regions for conceptualizing an action at different LOAs

Increasing LOAs were associated with the set of left hemisphere regions reliably associated with the *Why > How* contrast in previous studies (Figure 2; Spunt et al., 2010, 2011; Spunt and

Lieberman, 2012a,b; Spunt and Adolphs, 2014). These regions partially overlap with several meta-analytically defined functional networks, including: (i) the so-called 'theory-of-mind' or 'mentalizing' network associated with tasks of mental-state reasoning (Gallagher and Frith, 2003; Amodio and Frith, 2006; Saxe, 2006; Carrington and Bailey, 2009; Van Overwalle and Baetens, 2009; Schurz et al., 2014); (ii) the default mode network (DMN), especially its dorsomedial PFC component (Raichle et al., 2001; Buckner et al., 2008; Andrews-Hanna et al., 2010, 2014); (iii) the network associated with mentally simulating episodes both past and future (Hassabis and Maguire, 2007; Spreng et al., 2009; Schacter et al., 2012); (iv) the network associated with comprehending narrative discourse (Ferstl and von Cramon, 2001; Ferstl et al., 2008; Mar, 2011; Nijhof and Willems, 2015); (v) the network associated with transmodal semantic processing (Binder et al., 2009; Binder and Desai, 2011) and (vi) the network associated with comprehending abstract compared to concrete words (Binder et al., 2009; Wang et al., 2010; Binder and Desai, 2011).

Decreasing LOAs were associated with the majority of the left hemisphere regions reliably associated with the How > Why contrast in previous work. These regions fall within a broader functional network responsive to cognitive tasks devoid of meaningful socioemotional content (Van Overwalle, 2011; Jack et al., 2012). But they are more frequently regarded as forming a subset of the functional network thought to enable representation of the visual and somatomotor features of actions when they are perceived, performed, conceptualized, and verbally processed (Caspers et al., 2010; Kemmerer et al., 2012; Molenberghs et al., 2012; Pulvermuller, 2013; Watson et al., 2013; Rizzolatti et al., 2014; Urgesi et al., 2014; Kemmerer, 2015). This included a region of the left posterior MTG that has been associated with encoding the visual motion components of action concepts (Chen et al., 2008; Deen and McCarthy, 2010; Saygin et al., 2010; Wallentin et al., 2011; Humphreys et al., 2013; Watson et al., 2013). Interestingly, this seems to contrast with the view that the left posterior MTG represents more schematic aspects of the event structures encoded by both action and non-action verbs/sentences (Bedny et al., 2008, 2012).

Does the LOA construct help integrate this diverse set of findings? Several proposals have attempted to integrate at least a subset of these findings, but these have primarily regarded the regions associated with increasing LOA in the present study, which we have recently demonstrated map well on to the dorsomedial PFC subsystem of the DMN (Spunt et al., 2015). For instance, Buckner and Carroll (2007) suggest that mental-state reasoning, perspective-taking, episodic memory, and prospection all depend on a process they call 'self-projection', which involves the mental simulation of events and experiences that transcend the immediate environment. Mar and Oatley (2008) suggest a similar process of simulation to describe the intense social and emotional experiences that literary narratives are capable of evoking (see also Nijhof and Willems, 2015). Finally, Liberman and Trope (2014; Trope and Liberman, 2010) assert that mental representations can all be characterized by how 'psychologically distant' they are from 'a common zero-distance point, which is the experienced reality of me here and now' (Liberman and Trope, 2014, 1). While these theories have been primarily focused on the process of abstracting away from sensorimotor experiences (corresponding to increasing LOA in the present study), others have been primarily concerned with understanding the neural bases of relevant conceptual dualisms, such as between embodied and disembodied semantics (Pulvermuller, 2013) or between social and physical

knowledge domains (Van Overwalle and Baetens, 2009; Jack et al., 2012).

What these accounts all seem to share is a concern with understanding the human brain's ability for abstract thought, and distinguishing that ability from the one enabling concrete thought. The present study shares this concern in the important domain of action understanding, taking as its point of departure the idea that people naturally think about their actions as having a hierarchical structure. Such a point of departure overcomes a number of theoretical limitations of existing proposals. One such limitation is that prior proposals have typically focused on just one side of the conceptual coin, namely, on mental processes involved in abstraction. The present account parsimoniously handles both sides of the coin by identifying upward movements with abstraction and downward movements with the human ability to ground their abstract ideas in sensory experience and motor acts. Proposals that are concerned with both sides of the coin are limited by a tendency to frame the distinction as a conceptual dichotomy or dualism. The present account naturally permits representations to vary on a continuous dimension corresponding to the levels of a conceptual hierarchy. Finally, prior accounts have paid very little attention to identifying the specific task conditions and mental processes by which people shift their level of understanding. The present account is grounded in an ecologically valid method for eliciting such shifts: The Why/How Task.

Limitations and future directions

We identify several limitations of the present study that offer worthwhile directions for future research. First, we acknowledge the preliminary status of our empirical definition of LOA. The five semantic dimensions used to define the LOA factor were not intended to provide a complete and definitive list of semantic dimensions constituting LOA. There are likely many additional dimensions that could be included in an expanded definition of LOA. For instance, compare 'sharing knowledge' to 'typing words' as descriptions of the same act of writing a scientific paper. Compared to the latter, the former description gives the act a place in a pursuit that is more difficult, long-term, and socially relevant than does the latter description. Lin et al. (2014) recently showed that some of the brain regions tracking increasing LOA in the present study showed enhanced activation for action verbs that typically refer to social interactions (e.g., embrace), relative to verbs for individual actions (e.g. walk). They did not, however, measure any of the five dimensions of LOA featured in the present study, nor did they include RT as a nuisance covariate, making it difficult to ascertain to what extent a 'sociality' dimension adds anything distinctive to the concept of LOA as presented here. Similarly, it is likely that dimensions (e.g. mind vs body) that are useful for describing LOA in one conceptual domain (e.g. human action) will prove less useful for describing LOA in other conceptual domains (e.g. mathematical knowledge).

Second, we emphasize that we are not proposing that each of the five dimensions included in this study does not possess a unique and useful meaning on its own. The decision to focus on their shared variance was primarily motivated by the nature of our research question, which was to examine action conceptualization along a single dimension of hierarchical representation. Importantly, this decision was also empirically supported by a PCA on ratings of the five dimensions, which showed that a single factor could explain >91% of their total variance. Hence, we ultimately restricted our analyses to the derived LOA factor

and did not examine any of the dimensions individually. While these dimensions were highly correlated in the present study and may naturally correlate in language about actions, there is at least some evidence that they can be dissociated experimentally. For instance, recent work provides evidence for unique effects of concreteness, imageability, and valence in the neural responses to action verbs (Skipper and Olson, 2014; Vigliocco et al., 2014).

Third, we consider how our experimental task design could be adjusted to allow examination of novel questions regarding the neural bases of hierarchical action representation. In particular, we chose to present each question–answer pair in rapid succession, with the question offset and answer onset divided by a 250 ms blank screen (see Figure 1B). This had the benefit of minimizing working memory load and spontaneous answer production during the interval between question and answer presentation. This benefit came with a cost in that it prevented us from being able to separate the responses to question and answer presentation. Given that the magnitude of the Signed LOA Shift is ultimately established by presentation of the answer phrase, future studies may be able to more precisely model the onset of the within-trial shift by introducing an optimal amount of jitter between the question and answer onsets. However, modeling the onset of the shift at answer onset would still lack precision since the question marker (i.e. *why* vs *how*) provides unambiguous information about the sign of the imminent shift. This coupled with the high likelihood of spontaneous answer production in response to question presentation makes it likely that, in fact, the Signed LOA Shift of interest in our final model actually begins prior to answer presentation.

Finally, we identify two points of clarification regarding our claim that the primary function of *why* and *how* questions is to motivate increases or decreases in LOA, respectively. The first regards the possibility that, even when controlling for the stimulus content, the brain states evoked by *why* and *how* questions may be dissociable in subtle ways not detectable using univariate methods. Future studies could test for such effects using multivariate methods such as pattern-information analysis (Kriegeskorte and Kievit, 2013). The second regards the misinterpretation, briefly discussed above, that this means that *why* vs *how* is not to be considered a meaningful cognitive distinction. We believe the contribution of the present study is to more precisely specify the distinct effects that these two questions have on cognition. As discussed above, these effects require that the two questions are motivated by the same action stimulus. When this condition is met, it will almost always be the case that a *why* question will yield an answer that is at a higher LOA than will a *how* question. If this condition is not met, the effects will be unreliable and may sometimes reverse entirely. This is because, as noted in the Introduction, each question can be posed at any LOA, making it possible for a *why* question to yield an answer that rests at a lower LOA (Q: *Why* grip handlebars? A: Ride a bike.) than an answer to a *how* question (Q: *How* to stay healthy? A: Get exercise.). By way of summarizing, we offer the simple analogy: asking *why* vs *how* is akin to pressing the up vs down buttons when calling an elevator. As long as you're starting from the same floor, you can rest assured that pressing up will almost always put you on a higher floor than will pressing down.

Conclusion

We used a novel action understanding task with functional MRI to examine the neural basis of the well-documented effects of

answering *why* and *how* questions about actions. Our data conclusively demonstrate that these effects can be attributed to the fact that *why* and *how* questions—when asked of the same action—produce systematic changes in action understanding, and do so on what appears to be a single hierarchical dimension, which we refer to as the LOA (Vallacher and Wegner, 1985; Vallacher and Wegner, 1987). Increases and decreases in LOA tracked dissociable brain networks, consistent with prior work using the *Why/How* contrast. Our data particularly highlight the role of the dorsomedial PFC and anterior STS in upward shifts in LOA. Such shifts make it possible for people to conceive the ‘here-and-now’ of physical reality—including their own bodies—in abstract terms. This, in turn, gives us the power to appreciate that even the simplest of motor actions can carry information about who we are and what we care about most.

Acknowledgements

The authors would acknowledge Emily Ellsworth for help with data collection; Mike Tyszka and the Caltech Brain Imaging Center for help with the neuroimaging; and the Della Martin Foundation for postdoctoral fellowship support to R.P.S.

Funding

This work was supported in part by the National Institutes of Health (R01 MH080721-03 to R.A.).

Supplementary data

Supplementary data are available at SCAN online.

Conflict of interest. None declared.

References

- Amodio, D., Frith, C. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268–77.
- Andrews-Hanna, J., Reidler, J., Sepulcre, J., Poulin, R., Buckner, R. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron*, 65, 550–62.
- Andrews-Hanna, J., Smallwood, J., Spreng, R. (2014). The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences*, 1316, 29–52.
- Ashburner, J. (2007) A fast diffeomorphic image registration algorithm. *NeuroImage*, 38, 95–113.
- Badre, D., Poldrack, R., Pare-Blagoev, E., Insler, R., Wagner, A. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, 47, 907–18.
- Bedny, M., Caramazza, A., Grossman, E., Pascual-Leone, A., Saxe, R. (2008). Concepts are more than percepts: The case of action verbs. *Journal of Neuroscience*, 28, 11347–53.
- Bedny, M., Caramazza, A., Pascual-Leone, A., Saxe, R. (2012). Typical neural representations of action verbs develop without vision. *Cerebral Cortex*, 22, 286–93.
- Binder, J., Desai, R. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, 15, 527–36.
- Binder, J., Desai, R., Graves, W., Conant, L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19, 2767–96.

- Binder, J., Westbury, C., Mckiernan, K., Possing, E., Medler, D. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of Cognitive Neuroscience*, **17**, 905–17.
- Brainard, D.H. (1997). The psychophysics toolbox. *Spatial Vision*, **10**, 433–6.
- Buckner, R., Andrews-Hanna, J., Schacter, D. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, **1124**, 1–38.
- Buckner, R., Carroll, D. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, **11**, 49–57.
- Carrington, S., Bailey, A. (2009). Are there theory of mind regions in the brain? A review of the neuroimaging literature. *Human Brain Mapping*, **30**, 2313–35.
- Caspers, S., Zilles, K., Laird, A., Eickhoff, S. (2010). A meta-analysis of action observation and imitation in the human brain. *NeuroImage*, **50**, 1148–67.
- Chen, E., Widick, P., Chatterjee, A. (2008). Functional-anatomical organization of predicate metaphor processing. *Brain and Language*, **107**, 194–202.
- Deen, B., McCarthy, G. (2010). Reading about the actions of others: biological motion imagery and action congruency influence brain activity. *Neuropsychologia*, **48**, 1607–15.
- Denny, B., Kober, H., Wager, T., Ochsner, K. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *Journal of Cognitive Neuroscience*, **24**, 1742–52.
- Diedrichsen, J., Shadmehr, R. (2005). Detecting and adjusting for artifacts in fmri time series data. *NeuroImage*, **27**, 624–34.
- Ferstl, E.C., Neumann, J., Bogler, C., von Cramon, D.Y. (2008). The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mappings*, **29**, 581–93.
- Ferstl, E.C., von Cramon, D.Y. (2001). The role of coherence and cohesion in text comprehension: an event-related fMRI study. *Brain Research. Molecular Brain Research*, **11**, 325–40.
- Freitas, A.L., Gollwitzer, P., Trope, Y. (2004). The influence of abstract and concrete mindsets on anticipating and guiding others' self-regulatory efforts. *Journal of Experimental Social Psychology*, **40**, 739–52.
- Fujita, K., Trope, Y., Liberman, N., Levin-Sagi, M. (2006). Construal levels and self-control. *Journal of Personality and Social Psychology*, **90**, 351–67.
- Gallagher, H.L., Frith, C.D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences*, **7**, 77–83.
- Gilead, M., Liberman, N., Maril, A. (2014). From mind to matter: neural correlates of abstract and concrete mindsets. *Social Cognitive and Affective Neuroscience*, **9**, 638–45.
- Goldberg, R., Perfetti, C., Fiez, J., Schneider, W. (2007). Selective retrieval of abstract semantic knowledge in left prefrontal cortex. *Journal of Neuroscience*, **27**, 3790–98.
- Gollwitzer, P.M., Heckhausen, H., Steller, B. (1990). Deliberative and implemental mind-sets: Cognitive tuning toward congruous thoughts and information. *J Personality and Social Psychology*, **59**, 1119–27.
- Grinband, J., Wager, T., Lindquist, M., Ferrera, V., Hirsch, J. (2008). Detection of time-varying signals in event-related fmri designs. *NeuroImage*, **43**, 509–20.
- Hassabis, D., Maguire, E. (2007). Deconstructing episodic memory with construction. *Trends in Cognitive Sciences*, **11**, 299–306.
- Hoffman, P., Jefferies, E., Lambon Ralph, M. (2010). Ventrolateral prefrontal cortex plays an executive regulation role in comprehension of abstract words: Convergent neuropsychological and repetitive tms evidence. *Journal of Neuroscience*, **30**, 15450–56.
- Humphreys, G.F., Newling, K., Jennings, C., Gennari, S.P. (2013). Motion and actions in language: semantic representations in occipito-temporal cortex. *Brain and Language*, **125**, 94–105.
- Jack, A., Dawson, A., Begany, K., et al. (2012). Fmri reveals reciprocal inhibition between social and physical cognitive domains. *NeuroImage*, **66C**, 385–401.
- Kemmerer, D. (2015). Visual and Motor Features of the Meanings of Action Verbs: A Cognitive Neuroscience Perspective. In: de Almeida, R.G., Manouilidou, C., editors. *Cognitive science perspectives on verb representation and processing*. New York: Springer. p. 189–212.
- Kemmerer, D., Rudrauf, D., Manzel, K., Tranel, D. (2012). Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. *Cortex*, **48**, 826–48.
- Kozak, M., Marsh, A., Wegner, D. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, **90**, 543–55.
- Kriegeskorte, N., Kievit, R. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, **17**, 401–12.
- Liberman, N., Trope, Y. (2014). Traversing psychological distance. *Trends in Cognitive Sciences*, **18**, 364–9.
- Lieberman, M. (2010). Social cognitive neuroscience. In: Fiske, S.T., Gilbert, D.T., Lindzey, G., editors. *Handbook of Social Psychology*, 5th ed. New York: McGraw-Hill. p. 143–193.
- Lin, N., Bi, Y., Zhao, Y., Luo, C., Li, X. (2014). The theory-of-mind network in support of action verb comprehension: Evidence from an fMRI study. *Brain and Language*, **141C**, 1–10.
- Mar, R. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology*, **62**, 103–134.
- Mar, R.A., Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, **3**, 173–92.
- Molenberghs, P., Cunnington, R., Mattingley, J. (2012). Brain regions with mirror properties: A meta-analysis of 125 human fmri studies. *Neuroscience and Biobehavioral Reviews*, **36**, 341–49.
- Nijhof, A.D., Willems, R.M. (2015). Simulating fiction: individual differences in literature comprehension revealed with FMRI. *PLoS One*, **10**, e0116492.
- Pulvermuller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences*, **17**, 458–70.
- Raichle, M., Macleod, A., Snyder, A., Powers, W., Gusnard, D., Shulman, G. (2001). A default mode of brain function. *Proceedings of National Academy of Sciences U S A*, **98**, 676–82.
- Raposo, A., Mendes, M., Marques, J. (2012). The hierarchical organization of semantic memory: Executive function in the processing of superordinate concepts. *NeuroImage*, **59**, 1870–78.
- Rizzolatti, G., Cattaneo, L., Fabbri-Destro, M., Rozzi, S. (2014). Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiological Reviews*, **94**, 655–706.
- Satpute, A., Badre, D., Ochsner, K. (2014). Distinct regions of prefrontal cortex are associated with the controlled retrieval and selection of social information. *Cerebral Cortex*, **24**, 1269–77.
- Saxe, R. (2006). Uniquely human social cognition. *Current Opinion in Neurobiology*, **16**, 235–39.
- Saygin, A.P., McCullough, S., Alac, M., Emmorey, K. (2010). Modulation of BOLD response in motion-sensitive lateral

- temporal cortex by real and fictive motion sentences. *Journal of Cognitive Neuroscience*, 22, 2480–90.
- Schacter, D., Addis, D., Hassabis, D., Martin, V., Spreng, R., Szpunar, K. (2012). The future of memory: Remembering, imagining, and the brain. *Neuron*, 76, 677–94.
- Schurz, M., Radua, J., Aichhorn, M., Richlan, F., Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews*, 42C, 9–34.
- Shrout, P.E., Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychol Bulletin*, 86, 420.
- Skipper, L., Olson, I. (2014). Semantic memory: Distinct neural representations for abstractness and valence. *Brain and Language*, 130, 1–10.
- Spreng, R., Mar, R., Kim, A. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. *Journal of Cognitive Neuroscience*, 21, 489–510.
- Spunt, R., Adolphs, R. (2014). Validating the why/how contrast for functional mri studies of theory of mind. *Neuroimage*, 99, 301–11.
- Spunt, R., Falk, E., Lieberman, M. (2010). Dissociable neural systems support retrieval of how and why action knowledge. *Psychological Sciences*, 21, 1593–8.
- Spunt, R., Lieberman, M. (2012a). An integrative model of the neural systems supporting the comprehension of observed emotional behavior. *NeuroImage*, 59, 3050–9.
- Spunt, R., Lieberman, M. (2012b). Dissociating modality-specific and supramodal neural systems for action understanding. *Journal of Neurosciences*, 32, 3575–83.
- Spunt, R.P., Meyer, M.L., Lieberman, M.D. (2015). The default mode of human brain function primes the intentional stance. *Journal of Cognitive Neuroscience*, 27, 1116–24.
- Spunt, R., Satpute, A., Lieberman, M. (2011). Identifying the what, why, and how of an observed action: An fmri study of mentalizing and mechanizing during action observation. *Journal of Cognitive Neuroscience*, 23, 63–74.
- Trope, Y., Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Reviews*, 117, 440–63.
- Urgesi, C., Candidi, M., Avenanti, A. (2014). Neuroanatomical substrates of action perception and understanding: an anatomic likelihood estimation meta-analysis of lesion-symptom mapping studies in brain injured patients. *Frontiers in Human Neurosciences*, 8, 344.
- Vallacher, R.R., Wegner, D.M. (1985). *A Theory of Action Identification*. New York, NY: Psychology Press.
- Vallacher, R.R., Wegner, D.M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Reviews*, 94, 3–15.
- Van Overwalle, F. (2011). A dissociation between social mentalizing and general reasoning. *NeuroImage*, 54, 1589–99.
- Van Overwalle, F., Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage*, 48, 564–84.
- Vigliocco, G., Kousta, S., Della Rosa, P., et al. (2014). The neural representation of abstract words: The role of emotion. *Cerebral Cortex*, 24, 1767–77.
- Wallentin, M., Nielsen, A.H., Vuust, P., Dohn, A., Roepstorff, A., Lund, T.E. (2011). Bold response to motion verbs in left posterior middle temporal gyrus during story comprehension. *Brain and Language*, 119, 221–5.
- Wang, J., Conder, J., Blitzer, D., Shinkareva, S. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human and Brain Mappings*, 31, 1459–68.
- Watson, C.E., Cardillo, E.R., Ianni, G.R., Chatterjee, A. (2013). Action concepts in the brain: an activation likelihood estimation meta-analysis. *Journal of Cognitive Neuroscience*, 25, 1191–205.